JB0600Q

EXPRESSION OF SOLUBLE HETEROLOGOUS PROTEINS IN BACTERIA UTILIZING A THIOREDOXIN/PROTEIN EXPRESSION VECTOR

This application claims the benefit of U.S. Provisional Application 60/011,606, filed April 30, 1996.

BACKGROUND OF THE INVENTION

Expression of heterologous proteins in bacteria such as *E. coli* usually results in the formation of insoluble inclusion bodies that must be denatured and properly folded before the "natural" protein product is finally obtained. Thus there is a need to develop a bacterial expression system in which heterologous proteins can be expressed in the bacteria in a soluble, biologically active form. .

SUMMARY OF THE INVENTION

The present invention fills this need by providing for a vector which coexpresses a heterologous protein and thioredoxin wherein the heterologous protein and the thioredoxin are expressed as separate, non-fused proteins.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows schematically the construction of plasmid pDR75.
Figure 2 shows schematically the construction of plasmid pDR85.
Figure 3 shows schematically the construction of plasmid pDR109.
Figure 4 shows schematically the construction of plasmid pDR88.
Figure 5 shows schematically the construction of plasmid pDR80.
Figure 6 shows schematically the construction of plasmid pDR102.
Figure 7 shows schematically the construction of plasmid pDR112.

DETAILED DESCRIPTION OF THE INVENTION

All references cited herein are incorporated herein by reference.

According to the process of the present invention heterologous proteins can be produced properly folded, soluble and biologically active

by the coexpression of thioredoxin and the heterologous protein in bacteria especially *Escherichia coli* (*E. coli*). However, according to the present invention, the thioredoxin and the heterologous protein must be coexpressed as separate proteins and not as fused proteins.

As used herein, the term "transformed bacteria" means bacteria that have been genetically engineered to produce a mammalian protein. Such genetic engineering usually entails the introduction of an expression vector into a bacterium. The expression vector is capable of autonomous replication and protein expression relative to genes in the bacterial genome. Construction of bacterial expression is well known in the art, provided the nucleotide sequence encoding a desired protein is known or otherwise available. For example, DeBoer in U.S. Pat. No. 4,551,433 discloses promoters for use in bacterial expression vectors; Goeddel *et al.* in U.S. Pat. No. 4,601,980 and Riggs, in U.S. Pat. No. 4,431,739 disclose the production of mammalian proteins by *E. coli* expression systems; and Riggs *supra*, Ferretti *et al. Proc. Natl. Acad. Sci.83*:599 (1986), Sproat *et al., Nucleic Acid Research 13*:2959 (1985) and Mullenbach *et al., J. Biol. Chem 261*:719 (1986) disclose how to construct synthetic genes for expression in bacteria. Many bacterial expression vectors are available commercially and through the American Type Culture Collection (ATCC), Rockville, Maryland.

In the present invention a bacterium is transformed with vector containing a gene encoding a heterologous protein and a gene encoding a thioredoxin protein. An example of such a thioredoxin gene is SEQ ID NO:3. The following examples illustrate the coexpression of thioredoxin and heterologous proteins to produce properly folded proteins. The nucleic acid or gene which encodes the thioredoxin and the nucleic acid or gene which encodes the heterologous protein should be on the same vector such as a plasmid. Furthermore, it is even more preferable that the nucleic acid or gene which encodes the thioredoxin and the nucleic acid or gene which encodes the heterologous protein should be operationally linked to a common promoter such as the *lac* promoter.

Example 1

PCR Cloning of the Thioredoxin Gene from *E. coli*

E. coli chromosomal DNA was isolated from host strain MM294 according to the BioRad Instagene procedure. PCR primers were

synthesized according to the published sequence for the thioredoxin (trxA) gene. The forward primer includes an NdeI site within the methionine start codon such that the trxA gene may be readily cloned and expressed by the cytoplasmic pMBD vectors illustrated in the figures shown below. The reverse primer includes a silent nucleotide change to generate a BsaBI site for future constructions and a BamHI site for expression vector cloning.

Forward Primer (SEQ ID NO:1)

NdeI

CCTGTGGAGT TA**CATATG**AG CGATAAAATT

Reverse Primer (SEQ ID NO:2)

BamHI BsaBI

GCACCCAACA TGCAA**GGATC C**TTACGCCA**G ATTAGCATC**G AGGAACT

This resulted in the following trxA gene (SEQ ID NO:3)

ATGAGCGATA AAATTATTCA CCTGACTGAC GACAGTTTTG ACACGGATGT
ACTCAAAGCG GACGGGGCGA TCCTCGTCGA TTTCTGGGCA GAGTGGTGCG
GTCCGTGCAA AATGATCGCC CCGATTCTGG ATGAAATCGC TGACGAATAT
CAGGGCAAAC TGACCGTTGC AAAACTGAAC ATCGATCAAA ACCCTGGCAC
TGCGCCGAAA TATGGCATCC GTGGTATCCC GACTCTGCTG CTGTTCAAAA
ACGGTGAAGT GGCGGCAACC AAAGTGGGTG CACTGTCTAA AGGTCAGTTG
AAAGAGTTCC TCGATGCTAA TCTGGCGTAA GGATCC

A PCR product of the anticipated size was obtained, NdeI/BamHI digested and cloned into NdeI/BamHI digested pMBD202020 as outlined in the figures. The insert DNA was verified to be correct by nucleotide sequence analysis and the clone was designated pDR75-11. (Figure 1)

Example 2

Construction of a regulated vector containing the trxA vector

Vector pDR75-11 is a constitutive expression vector and it was desired to have a vector in which the expression of the trxA gene could be regulated. The trxA gene from pDR75-11 was subcloned as a XbaI/BamHI fragment into pMBD112012. The resulting plasmid was

designated pDR85. The trxA gene is expressed from the lpp/lac promoter-operator and is regulated by the lacIQ repressor. (Figure 2)

Example 3

Plasmid pDR109 Construction (Figure 3)

The trx A gene was altered to include a unique XhoI restriction site to allow for easy subcloning of a downstream recombinant protein. The trxA gene was PCR amplified.

A forward primer incorporated four nucleotide changes from the wild type *E. coli* DNA sequence so as to optimize the codon usage within the first five codons because optimal codon usage has been known to increase the efficiency of translation initiation. A reverse primer includes the incorporation of the XhoI site which results in a conservative amino acid change (aspartate to glutamate) in the thioredoxin protein.

The PCR product was subcloned into pMBD112012. The resulting plasmid expresses thioredoxin as a cytoplasmic protein from the $lacI^Q$ regulated lpp-lac promoter on a pBR322 replicon.

Shown below is the resultant trxA gene in pDR109 (SEQ ID NO:4)

```
ATGAGCGATA AAATTATTCA CCTGACTGAC GACAGTTTTG ACACGGATGT
ACTCAAAGCG GACGGGGCGA TCCTCGTCGA TTTCTGGGCA GAGTGGTGCG
GTCCGTGCAA AATGATCGCC CCGATTCTGG ATGAAATCGC TGACGAATAT
CAGGGCAAAC TGACCGTTGC AAAACTGAAC ATCGATCAAA ACCCTGGCAC
TGCGCCGAAA TATGGCATCC GTGGTATCCC GACTCTGCTG CTGTTCAAAA
ACGGTGAAGT GGCGGCAACC AAAGTGGGTG CACTGTCTAA AGGTCAGTTG
AAAGAGTTCC TCGAGGCTAA TCTGGCGTAA GGATCC
```

Coexpression of thioredoxin and the recombinant protein is achieved by mimicking the translational coupling which occurs naturally in the tryptophan operon of *E. coli*. The ribosome binding site for the downstream gene is located within the 3' end of the preceding coding region and the stop and start codons of the adjacent genes are either overlapping or are immediately adjacent to each other.

The translationally coupled recombinant gene is generated by PCR amplification with a forward primer which includes the XhoI cloning site, sequences for the ribosomes binding site within the 3' end

of the trxA gene, the stop codon for trxA (TAA) and the ATG start codon and the beginning DNA nucleotides of the recombinant gene. The incorporation of the ribosome binding site sequences within the 3' end of the trxA gene results in non-conservative amino acid changes within the protein.

Example 4

Construction of a trxA/recombinant Human Interleukin-13

Vector pDR88 contains the trxA/recombinant human IL-13 (rhuIL-13) gene fusion with a gly/ser hinge linker + enterokinase cleavage site as described by LaVallie, *et al.* (Figure 4)

Linkers were attached to a rhuIL-13 clone (pLET3) which generated pDR80. The linkers contain the BsaBI site + gly/ser hinge linker + enterokinase cleavage site + rhu IL-13 codons + SstI site. (Figure 5)

The BsaBI/BAMHI fragment from pDR80 was cloned into pDR85 to generate pDR88. (Figure 6)

Sequence of the U411/U412 linker region ~~(SEQ ID NO:5)~~

BsaBI

```
GAT AAT ATT CTG GCT GGT TCT GGT TCT GGT GAT GAC GAT GAC AAG
Asp Asn Asn Leu Ala Gly Ser Gly Ser Gly Asp Asp Asp Asp Lys
---trxA-----||Gly/Ser hinge ----||enterokinase cleavage

                                                  SstI
GGT CCT GTT CCG CCG TCT ACC GCT CTG CGT GAG CTC
Gly Pro Val Pro Pro Ser Thr Ala Leu Arg Glu Leu
```

Example 5

trx Gene Translationally coupled to the rhuIL-13 Gene

A BsaBI/Sst linker was synthesized to include a ribosome binding site and coupled stop/start codon for trxA/rhu IL-13. The double stranded oligo was cloned into pDR88 to generate pDR102. (Figure 6)

Translational Coupling Sequence in pDR102 (SEQ ID NO:7) ~~(SEQ ID NO:6)~~

```
R.B.S.          IL-13                    Sst I
```

GAAGGAGGCT GATTAA**ATGGGTCCGGTTCCGCCGTCTACCGCTCTGGAGCTC**

Recombinant Human IL-13 (rhu IL-13) was translationally coupled to thioredoxin with the following sequence: (SEQ ID NO:7)

```
    RBS    |----8bp---|
------AAG GAG GCT GAT TAA ATG----------
--trxA-----------------|Met ---rhuIL-13
```

The resultant plasmid (designated pDR102) (Figure 6) was transformed into *E. coli* host strain MM294 and fermentation analysis was done to confirm protein expression. The culture was induced for expression and grown at 15°C to maximize the accumulation of rhuIL-13 soluble protein. Cells were harvested at 48 and 68 hrs post induction. Accumulation of soluble protein immunoreactive to anti-IL-13 antibody and of monomeric (non-fused) size was observed at both times.

Example 6

Fermentation Analysis of rhu IL-10 and rhuIL-13 Production from trxA Plasmids

Alternative coupling sequences were analyzed for rhuIL-13 clones. The two alternative sequences in pDR113 and pDR114 differ from pDR102 in that the stop codon (TAA) for trxA and the start codon (ATG) for rhIL-13 overlap each other as the TAATG sequence. In addition, the spacing between the ribosome binding site (RBS) and the ATG start codon is shorter, reduced to 7 bp in pDR113 and to 4 bp in pDR114.

```
     RBS   |--7bp--|
-----AAG GAG GCT GAT TAATG----  (SEQ ID NO:8) pDR113
--trxA-----------     |Met--rhuIL-13
      RBS    |-4 bp-|
------ AAG GAG GTT TAATG---
------trxA-----------|Met---rhu IL-13 (SEQ ID NO:9)    pDR114
```

Fermentations were done at 15°C. Soluble protein is produced in pDR113 and pDR114.

Attempts were made to enhance protein expression from pDR102 by using the Tac promoter instead of the lpp-lac promoter and by

increasing plasmid copy number by utilizing the pUC origin of replication.

Plasmid pDR111 contains the pDR102 coupling expressed from the Tac promoter. Plasmid pDR112 utilizes the pDR102 coupling expressed from the Tac promoter and pUC origin of replication. (Figure 7)

Fermentations were done at 15°C. Soluble protein was produced in both pDR111 and pDR112.

Example 7

Coexpression of Thioredoxin and Recombinant Human Interleukin-10

A trxA/rhuIL-10 fusion plasmid was made and designated pDR130. Fermentations were performed at 15°C, 25°C and 37°C. Production of soluble trxA-rhuIL-10 fusion protein was greatest at 15°C and still detectable at 37°C. Protein material remained in the soluble fraction after 90 minutes centrifugation at 40,000 rpm.

Many modifications and variations of this invention can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. The specific embodiments described herein are offered by way of example only, and the invention is to be limited only by the terms of the appended claims, along with the full scope of equivalents to which such claims are entitled.